DOCUMENT SUMMARIZATION BASED ON TOPICALITY AND SPECIFICITY

Abstract of the Disclosure

5   Topicality scores are determined for a number of phrasal expressions in documents. Phrasal expressions may be noun phrases, with or without corresponding prepositional phrases, subject-verb pairs, and verb-object pairs. The documents describe some topic or multiple topics. Techniques can be used to determined how the phrasal expression compares with the topic or topics being described in the documents. 10 Specificities are determined for the phrasal expressions. Techniques may be used to determine whether phrasal expressions are more or less specific than other phrasal expressions. An order is determined for the phrasal expressions by using the topicality scores and the specificities. The order may be represented as a phrasal expression tree, for example. The phrasal expression tree may be displayed to a user, and the user can 15 navigate through the phrasal expression tree, and therefore through the one or more documents.